HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2016 May 18.

Published in final edited form as:

Nat Neurosci. 2015 December ; 18(12): 1804–1810. doi:10.1038/nn.4158.

Inferring learning rules from distribution of firing rates in cortical neurons

Sukbin Lim[1], **Jillian L. McKee**[1], **Luke Woloszyn**[2], **Yali Amit**[3,4], **David J. Freedman**[1], **David L. Sheinberg**[5], and **Nicolas Brunel**[1,3,*]

[1]Department of Neurobiology, University of Chicago, Chicago, IL 60637, USA

[2]Department of Neuroscience, Columbia University, New York, NY 10032, USA

[3]Department of Statistics, University of Chicago, Chicago, IL 60637, USA

[4]Department of Computer Science, University of Chicago, Chicago, IL 60637, USA

[5]Department of Neuroscience, Brown University, Providence, RI 02912, USA

Abstract

Information about external stimuli is thought to be stored in cortical circuits through experience-dependent modifications of synaptic connectivity. These modifications of network connectivity should lead to changes in neuronal activity, as a particular stimulus is repeatedly encountered. Here, we ask what plasticity rules are consistent with the differences in the statistics of the visual response to novel and familiar stimuli in inferior temporal cortex, an area underlying visual object recognition. We introduce a method that allows inferring the dependence of the 'learning rule' on post-synaptic firing rate, and show that the inferred learning rule exhibits depression for low post-synaptic rates and potentiation for high rates. The threshold separating depression from potentiation is strongly correlated with both mean and standard deviation of the firing rate distribution. Finally, we show that network models implementing a rule extracted from data show stable learning dynamics, and lead to sparser representations of stimuli.

Introduction

Reorganization of neuronal circuits through experience-dependent synaptic modification has been postulated to be one of the basic mechanisms for learning and memory[1]. This idea is supported by experimental work from different preparations that show long-term changes of synaptic strengths induced by various patterns of pre- and post-synaptic activity[2–6]. Such activity-dependent synaptic modifications in a neural circuit would in turn lead to changes of activity of the circuit. A positive feedback between synaptic potentiation and elevated

*Corresponding author: ; Email: nbrunel@uchicago.edu

Author Contributions S.L., Y.A. and N.B. designed the research. S.L. analyzed the data, performed network simulations and prepared the figures. J.L.M., L.W., D.J.F. and D.L.S. contributed the electrophysiological data. S.L. and N.B. wrote the manuscript, with contributions from all authors.

The authors declare no competing financial interests.

neuronal activity could lead to enhanced neuronal responses, while synaptic depression would lead to opposite changes.

While changes of synaptic strengths in strongly connected cortical circuits are difficult to identify *in vivo*, changes of single neuron responses *in vivo* have been suggested as evidence for synaptic plasticity in cortical circuits. In particular, perturbations in input statistics or perceptual learning tasks have been shown to induce changes in neuronal responses[7–9]. Theoretical models have been used to understand interactions between activity-dependent plasticity rules and network activity[10,11]. Such models typically implement synaptic plasticity rules extracted from *in vitro* studies, and provide qualitative explanations for changes of sensory representations in feed-forward circuits[12,13], and changes of sensory and memory-related activity in recurrently connected circuits[14,15].

One of the cortical areas where the effects of sensory experience on neuronal responses have been documented is inferior temporal cortex (ITC), an area which is critical for visual object perception and recognition[16–18]. Two types of experiments have been used, one in which initially novel visual stimuli are shown repeatedly to a monkey[19], and another where two sets of stimuli (novel and familiar) are compared[18,20–24]. Several effects of visual experience on ITC neuronal activity and selectivity have been described in these studies. First, it has been shown that repeated presentations of an initially novel stimulus in a single recording session, leads to a gradual decrease of visual responses to the stimulus in a significant fraction of recorded neurons[19]. Second, a comparison between visual responses to novel stimuli and stimuli that have been presented over many recording sessions have demonstrated that the response to familiar stimuli is typically more selective[18,20–23], with higher maximum responses to familiar stimuli in putative excitatory neurons[22]. However, it is still unclear what type of learning rules could explain this data.

Here, we introduce a procedure that allows us to derive the synaptic plasticity rule from changes in distributions of visual responses to novel and familiar stimuli, using a cortical network model composed of excitatory and inhibitory neurons, whose excitatory-to-excitatory connectivity is plastic. We applied this method to experimental data obtained in ITC neurons in monkeys performing two different tasks, a passive-fixation task[22], and a dimming-detection task[25]. Finally, we showed that simulations implementing learning rules derived from data in a recurrent network model provides a good match with experimental data.

Results

Changes in network response induced by synaptic plasticity

To investigate the relation between synaptic plasticity rule and changes of network activity with learning, we considered a firing rate model with a plasticity rule that modifies the strength of recurrent synapses as a function of the firing rates of pre- and post-synaptic neurons. Activities of neurons are described by their firing rates r_i, i=1,...N, where N denotes the number of neurons in the network. The firing rate of neuron i depends on its inputs h_i, through a static transfer function (f-I curve) Φ_i as

$$r_i = \Phi_i(h_i) \quad \text{with} \quad h_i = I_{iX} + \sum_{j=1}^{N} W_{ij} r_j. \tag{1}$$

The input current h_i is the sum of the external input I_{iX} and the recurrent input, which is itself a sum of pre-synaptic firing rates r_j, weighted by the synaptic strength W_{ij} connecting neuron j to neuron i.

We investigated how the network response changes with visual experience as initially novel stimuli become familiar. Changes due to visual experience could in principle come from changes in external inputs, recurrent inputs, or both. Here, we assume that changes in network response are primarily due to changes in recurrent synapses. This assumption is justified by the observation that differences between responses to familiar and novel stimuli start to emerge a few tens of milliseconds after the activity onset[21-23] (Fig. 1). We therefore assumed that the recurrent synapses are plastic, changing their strength according to $W_{ij} \rightarrow W_{ij} + \Delta W(r_i, r_j)$, where the synaptic change $\Delta W(r_i, r_j)$ depends on firing rates of both pre- and post-synaptic neurons during the presentation of the stimulus. The changes in synaptic strengths lead to changes in synaptic inputs to neurons, $h_i \rightarrow h_i + \Delta h_i$ and consequently to changes in their firing rates $r_i \rightarrow r_i + \Delta r_i$, according to

$$r_i + \Delta r_i = \Phi_i(h_i + \Delta h_i) \quad \text{with} \quad h_i + \Delta h_i = I_{iX} + \sum_{j=1}^{N} (W_{ij} + \Delta W(r_i, r_j))(r_j + \Delta r_j). \tag{2}$$

Note that since synaptic plasticity occurs only in recurrent connections, changes of recurrent synaptic inputs are the only source of input changes, so that henceforth input changes refer only to changes of recurrent inputs unless specified otherwise.

When $\Delta W(r_i, r_j)$ and Δr_j are small (compared to W_{ij} and r_j, respectively), $\sum_{j=1}^{N} \Delta W(r_i, r_j) \Delta r_j$ can be neglected in comparing Eq. (1) and Eq. (2), and the changes in inputs become approximately

$$\Delta h_i \sim \sum_{j=1}^{N} \Delta W(r_i, r_j) r_j + \sum_{j=1}^{N} W_{ij} \Delta r_j. \tag{3}$$

Eq. (3) shows two different contributions of $\Delta W(r_i, r_j)$ onto input changes Δh_i – the first term on the right-hand side represents a direct influence of changes of synaptic strengths projecting on a post-synaptic neuron on the inputs to that neuron, and the second term represents an indirect influence through changes of pre-synaptic firing rates that reflect changes of synaptic strengths in the whole network. Based on the above equations, we can calculate changes in firing rates Δr_j when the transfer functions Φ_i and the synaptic plasticity rule $\Delta W(r_i, r_j)$ are known (see the analytic expression of firing rates and its application to an example learning rule in Supplementary Math Note, section 1).

Inferring learning rules from firing rate distributions

We now ask the inverse question – can we infer the synaptic plasticity rule from the changes of network responses with learning? To perform this inference, we make the assumption that the learning rule is a separable function of pre and post-synaptic rates, $\Delta W(r_i, r_j) = f_{post}(r_i) f_{pre}(r_j)$. Then, the dependence of the learning rule on post-synaptic firing rates $f_{post}(r_i)$ can be obtained from Eq. (3) as

$$\begin{aligned}\Delta h_i &\sim \sum_{j=1}^{N} \Delta W(r_i, r_j)\, r_j \\ &+ \sum_{j=1}^{N} W_{ij} \Delta r_j \\ &= f_{post}(r_i) \sum_{j=1}^{N} f_{pre}(r_j)\, r_j \\ &+ \sum_{j=1}^{N} W_{ij} \Delta r_j \Rightarrow f_{post}(r_i) \sim \left(\Delta h_i - \sum_{j=1}^{N} W_{ij} \Delta r_j \right) / \sum_{j=1}^{N} f_{pre}(r_j)\, r_j. \end{aligned} \quad (4)$$

These equations allow us to estimate the function $f_{post}(r_i)$ characterizing how synaptic plasticity depends on the post-synaptic firing rate from the distributions of responses to novel and familiar stimuli, using a few additional assumptions that we detail below. The method is illustrated in Fig. 2 using the distributions of firing rates of a single ITC neuron, computed from the response of the neuron to 125 novel stimuli and 125 familiar stimuli[22] (Fig. 2a; see Online Methods). These distributions are similar to log-normal distributions, as has been noted previously in other areas[26–28]. The distribution of responses to familiar stimuli is shifted to the left compared to the distribution of responses to novel stimuli, indicating that on average, familiar stimuli elicit lower rates than novel stimuli, but the tail of the distribution of familiar responses extends further to the right compared to the distribution of novel responses.

From these empirical distributions, the inference of synaptic changes was done in four steps (Supplementary Fig. S1). First, we deduced the transfer function Φ from the distribution of responses for novel stimuli, under two assumptions: (i) the transfer function is monotonically increasing, that is, higher input currents lead to higher output firing rates, consistent with both *in vivo* and *in vitro* electrophysiological data[29,30], and (ii) input currents for novel stimuli follow a normal distribution. This assumption is based on the central limit theorem: Since cortical neurons receive a very large number of inputs which in the case of novel stimuli can be expected to be weakly correlated, the distribution of total synaptic inputs is likely to follow approximately Gaussian statistics. Under these two assumptions, the transfer function (*f-I* curve) is obtained from the empirical distribution of rates for novel stimuli and the assumed Gaussian distribution of input currents in the following procedure. Input currents and rates are ordered according to rank (Fig. 2b,c). Because the *f-I* curve is assumed to be monotonically increasing, ranks are preserved and the firing rate dependence on the input current is obtained by matching corresponding ranks (Fig. 2d). The obtained *f-I*

curve resembles qualitatively the experimentally measured *f-I* curves of pyramidal cells in the presence of noise[29,30].

Input currents for familiar stimuli were obtained from firing rates by applying the inverse of the derived transfer function, assuming that the transfer function remains unchanged with learning (blue curves in Fig. 2e,f). We then compare the distributions of input currents for novel and familiar stimuli, and extract how input changes with learning as a function of the response to novel stimuli. Since we do not have recordings of the neuron's response to the *same* stimulus as it transitions from novel to familiar, to perform this step, we need one last assumption: Given a set of initially novel stimuli leading to a set of corresponding responses, learning these stimuli does not change the rank of the corresponding responses. This allows us to compare responses to two different stimuli, one novel and one familiar, that have the same rank. For example, for a novel stimulus that elicits a median response (about 15 Hz in the example of Fig. 2e), we expect learning to decrease the rate to about 10 Hz, the median response to familiar stimuli. Under the rank preservation assumption, the input changes are computed by comparing input currents for novel and familiar stimuli at the same rank (Fig. 2e,f). Input changes are plotted as a function of post-synaptic firing rate before learning, i.e., Δh as a function of r_{post} (see Eq. (3); Fig. 2g). We remark that the rank preservation assumption minimizes $\sum(\Delta r_i)^2$ among all possible sets of Δr_i, (Supplementary Math Note, section 2). Thus, it is consistent with the assumption of small changes of synaptic strengths and firing rates, ΔW_{ij} and Δr_j, in our derivation of learning rules in Eq. (3) and (4).

Finally, from the function $\Delta h(r_i)$, we obtained a synaptic plasticity rule whose dependence on the post-synaptic firing rate $f_{post}(r_i)$ has a similar form to that of the input changes (Fig. 2g). According to Eq. (4), $f_{post}(r_i)$ can be obtained by subtracting input changes due to changes of firing rates $\sum_j W_{ij}\Delta r_j$ from total input changes $\Delta h(r_i)$, and normalizing it with a term containing the dependence of the learning rule on pre-synaptic firing rates $\sum_j f_{pre}(r_j)\, r_j$. Note that $\sum_j f_{pre}(r_j)\, r_j$ is constant for all neurons, and we assume that $\sum_j W_{ij}\Delta r_j$ is approximately independent of r_i. Thus, the dependence of the learning rule on post-synaptic firing rate $f_{post}(r_i)$ can be obtained from input changes by subtracting a constant offset and rescaling its magnitude (green line and axis on the right side in Fig. 2g).

The derived learning rules depend on the assumptions we made on input statistics, properties of the transfer function, and properties of the learning rules. However, our results remain qualitatively unchanged if these assumptions are relaxed. Transfer functions and learning rules derived assuming non-Gaussian statistics of input currents for novel stimuli (Supplementary Fig. S2) or starting with the input statistics for familiar stimuli (Supplementary Fig. S3), are qualitatively similar to the original ones. Also, we note that even if the assumption of the rank preservation of a stimulus is relaxed with the addition of noise to input currents, it is found that input changes computed under the relaxed assumption show similar dependence on post-synaptic firing rates (Supplementary Fig. S4).

Inferred ITC learning rules

Using the method illustrated above, we investigated the effect of visual experience in inferior temporal cortex (ITC), using neurophysiological data obtained from two different laboratories in monkeys performing two different tasks – one is a passive viewing task[22] and another is a dimming-detection task[25] (see Online Methods). In both cases, we obtained the distributions of neural activities to novel or familiar stimuli by taking firing rates of neurons averaged during the stimulus presentation period.

Despite the difference in tasks and the number of stimuli used to measure activities in each neuron, we found similar input changes with learning in the two data sets whose shape depends on cell type – putative excitatory or putative inhibitory (Fig. 3). When the distributions of firing rates for each novel and familiar stimulus are averaged over all recorded neurons, input currents show negative changes with learning for all post-synaptic firing rates, which are below the 95% confidence region obtained from computing this curve using random samples from the novel distribution instead of the familiar distribution (grey area in Fig. 3a,d, see Online Methods). This is consistent with experimental observations that average firing rates decrease with familiarity[21–23]. Next, we applied the analysis separately to putative excitatory and inhibitory cells, defined by the width of action potential waveforms (see Online Methods). Excitatory neurons showed negative changes when the post-synaptic firing rate is low, but positive changes when it is high (Fig. 3b,e). Such positive input changes in a high firing rate regime are consistent with the increase of maximal responses of excitatory neurons with learning that have been observed experimentally[16,18,20,22]. Inhibitory neurons showed negative input changes at all firing rates (Fig. 3c,f). Note that averaging over both groups of neurons completely masks the enhancement of input currents in excitatory neurons at high rates (Fig. 3a,d), because of the stronger negative input changes in inhibitory neurons (Fig. 3c,f).

In the experimental data obtained during the passive viewing task, we further analyzed the learning effects on input currents in individual neurons (Fig. 4). To compare input changes in neurons with a different range of firing rates, we normalized firing rates of post-synaptic neurons by subtracting their mean and dividing by the standard deviation of firing rates to different stimuli (Fig. 4a,b). In excitatory neurons, we found diverse patterns of input changes that can be classified into 3 categories (Fig. 4a) – neurons showing only negative changes (orange curves), neurons showing negative changes for low firing rates and positive changes for high firing rates (green curves), and neurons showing only positive changes (blue curves). Despite diverse patterns, the input changes in the excitatory neurons are increasing as normalized firing rates increase, except for a few neurons showing only negative changes and having high mean firing rates (Fig. 4a). These diverse patterns of input changes may result from a different constant offset in input changes in each neuron (the first term in the right hand side of Eq. (3)), rather than different forms of synaptic plasticity rules onto different neurons. Averaging the input change curves of excitatory neurons showing both negative and positive changes (green curves in Fig. 4a) leads to depression for low firing rates and potentiation for high firing rates, consistent with previous *in vitro* experiments[6,31] (Fig. 4c). Note also that the transfer functions of different neurons with normalized firing rates are very similar to each other (Supplementary Fig. S5).

For neurons showing both negative and positive changes, we defined a threshold θ as the post-synaptic firing rate where input changes become positive, and denote the normalized threshold by θ', obtained by subtracting the mean rate and dividing by the standard deviation of the rate (Fig. 4a). We found that the threshold θ is strongly correlated with both mean and standard deviation of post-synaptic firing rates (Fig. 4e,f), but such correlation disappears for the normalized threshold θ' (Fig. 4g,h). This suggests that a threshold between depression and potentiation in the synaptic plasticity rule is dependent on activity of neurons, which is reminiscent of the Bienenstock-Cooper-Munro learning rule[12]. The threshold observed in ITC neurons is around 1.5 standard deviations above the mean firing rate (Fig. 4c), consistent with a scenario in which a large majority of stimuli lead to depression, while a small minority (the ones with the strongest responses) lead to potentiation.

Changes of input currents derived from the statistics of individual inhibitory neurons were similar to those from the statistics of their population average – except for one neuron showing positive changes, inhibitory neurons show negative input changes (Fig. 4b) that depend only weakly on firing rates (Fig. 4d). Firing rate-independent negative changes can be explained by a decrease in average firing rates of the excitatory sub-network that would cause decrease in inputs from excitatory to inhibitory neurons (see Eq. (3)), in the absence of any plasticity mechanism in inhibitory neurons. A similar shape of negative input changes was observed in a few putative excitatory neurons (4 orange curves showing only decrease in Fig. 4a). These neurons have similar dynamic range of activity as inhibitory neurons with high mean firing rates, and may be inhibitory neurons with broad spike widths.

Simulations and comparison to the data

We next addressed whether a network model with a learning rule inferred from data can maintain stable learning dynamics as it is subjected to multiple novel stimuli, and whether the changes of activity patterns with learning observed in the experiment can be reproduced.

Simulated networks were composed of excitatory (E) and inhibitory (I) neurons and were initially fully connected. All neurons of a given type (E or I) had the same input-output transfer functions. Cells of the same type had the same distributions of firing rates, which were derived from data (see Online methods). Synaptic plasticity was implemented in excitatory-to-excitatory (E-to-E) connections only, while all connections involving inhibition were fixed. The motivation for restricting plasticity to E-to-E connections is that, as we will see below, this is the simplest, yet biologically plausible model that is able to reproduce the data (see also Supplementary Math Note, section 3). All E-to-E synapses had the same firing rate-dependent learning rule whose post-synaptic dependence was derived from data, taking into account excitatory neurons showing depression at low rates and potentiation at high rates (green curve in Fig. 4c). For simplicity, we took the dependence on the pre-synaptic rates f_{pre} to be linear (see below). We also added bounds on synaptic strengths. Note that in simulations, the synaptic strengths were updated only once per stimulus. This single update of synaptic strengths per stimulus effectively captures changes of synaptic strengths that occur gradually in time through multiple presentations, until the steady state of learning is reached. Synapses from or onto inhibitory neurons were assumed to be uniform with fixed strengths.

We initialized the excitatory connections by presenting a large number of uncorrelated activity patterns sampled from the distributions of firing rates to novel stimuli until a stable state was reached, and investigated learning dynamics at this stable state. Note that the input changes in the data showed depression for most of the post-synaptic firing rates. When we assumed that the dependence of the learning rule on the pre-synaptic firing rates is $f_{pre}(r_j) = r_j$ in Eq. (4), the total sum of synaptic weights $\sum_{i,j} f_{post}(r_i) f_{pre}(r_j)$ decreased with learning. Thus, synaptic weights were stabilized after learning multiple novel stimuli without encountering instabilities that are typical of Hebbian plasticity rules[32,33]. However, such a learning rule led to very low mean weights, since excitatory synaptic weights cannot be negative (Supplementary Fig. S6). As a consequence, the effects of learning one particular stimulus on responses were much smaller than the experimentally observed ones. To prevent this, we introduced a constraint under which the total sum of synaptic strengths onto post-synaptic neurons is preserved[34,35] (see Online Methods). This constraint is equivalent to $f_{pre}(r_j) = r_j - \text{mean}(r)$ where mean(r) is the population average of firing rates of pre-synaptic neurons. Such a constraint keeps $f_{post}(r_i)$ unchanged (see Online methods and Supplementary Math Note, section 4).

After the initialization stage in which many patterns were presented to the network, we compared the responses of all neurons in the network to a given stimulus before and after learning. Since the network is homogenous, and since novel firing rates for stimuli in the simulated network are sampled independently from the empirical distribution of novel firing rates, this can serve as a surrogate for comparing activities measured in a single neuron in response to multiple stimuli. Activity patterns before learning (red curves in Fig. 5c,d) were sampled from the distribution of firing rates to novel stimuli obtained from experiment (grey curves in Fig. 5a,b). The distributions of the firing rates after learning (blue curves in Fig. 5c,d) were obtained from simulating network dynamics with updated excitatory synaptic weights. They were similar to those from the experimental data (black curves in Fig. 5a,b). Average responses were reduced for both excitatory and inhibitory neurons, while a small fraction of excitatory neurons with high firing rates showed an increase (Fig. 5c,d). Consistently, most percentiles of the firing rate distribution decreased with learning in excitatory and inhibitory neurons (Fig. 5e,f), except for firing rates of excitatory neurons at high percentiles (e.g., the 95th percentile in Fig. 5e). These changes lead to increased selectivity and increased sparseness for learned stimuli, as observed in the experimental data[21,22].

Because synapses onto inhibitory neurons were not plastic, changes of firing rates in inhibitory neurons reflected changes of the average firing rate of excitatory neurons. A decrease in average firing rate of excitatory neurons led to a decrease in rates for all stimuli. Hence, despite its simplicity, network simulations implementing the learning rule show stable learning dynamics and reproduce quantitatively the main features observed in the data, including the decrease of average activities in both excitatory and inhibitory neurons and the increase of selectivity in excitatory neurons.

Discussion

We have introduced a method to derive input changes and learning rules from changes of firing rates in cortical neurons. In a cortical network model in which the learning rules are separable functions of pre- and post-synaptic rates, the inferred transfer functions were consistent with *in vivo* and *in vitro* electrophysiological data[29,30], and the dependence of the excitatory synaptic plasticity rule on post-synaptic firing rate was consistent with *in vitro* synaptic plasticity data[6,31]. Application of this method to experimental data from ITC neurons revealed several features of the inferred learning rules. First, the inferred learning rule in recurrent excitatory connections exhibits depression for low post-synaptic firing rates and potentiation for high rates, with a dominant effect of depression on average. Such a depression-dominant learning rule leads to the decrease of average firing rates observed in the data, which cannot be captured by previously suggested Hebbian learning rules[36,37]. Second, the threshold separating depression from potentiation is strongly correlated with the mean and standard deviation of post-synaptic firing rates, which suggests a regulation of learning rules depending on neuronal activity. Third, experimentally observed changes in inhibitory firing patterns can be induced by changes of excitatory firing patterns without synaptic changes onto inhibitory neurons, although the data does not rule out such changes. Finally, implementation of the inferred learning rule in a recurrent network model shows stable learning dynamics and provides a good match to the data.

One of the key features of the inferred learning rule is a strong correlation between the threshold separating depression and potentiation and neural activity. This is reminiscent of the Bienenstock-Cooper-Munro learning rule where the threshold separating depression and potentiation is a dynamic function of the post-synaptic firing rate, leading to stabilization of learning[12,38]. However, it has been noted that the time scale for the sliding threshold needs to be sufficiently fast to avoid such instabilities[32,33]. Our data does not allow us to investigate a possible dynamic evolution of the threshold and its contribution to stable learning, since it gives us only a ‘snapshot’ in the life of neurons in ITC. In addition to its role in stabilizing learning dynamics, such a regularization of the plasticity rule may contribute to enhancing selectivity to stimuli in heterogeneous populations – given that neurons have different ranges of firing rates, an adjustment of threshold depending on neuronal activity enables potentiation and depression to occur in all neurons, leading to an expansion of the range of population responses and an enhancement of selectivity with learning.

Our work provides a method to infer synaptic plasticity rules from responses to two large sets of stimuli, one novel, the other familiar. It could also be applied to an experiment that traces responses to initially novel stimuli as the stimuli became familiar. Such an experiment would in addition reveal the speed of learning, potentially provide information on an eventual dynamic evolution of the threshold separating depression and potentiation, and test the rank preservation assumption. Similar experiments have been performed previously[19], but responses to a sufficiently large number of stimuli would need to be measured, in order to access the potentiation region of the input change vs firing rate curve. We predict that in such an experiment, the stimuli that elicit initially the highest firing rates would see an

increase in firing rates as they are presented repeatedly, provided a sufficiently large number of stimuli is used[22].

While our model shows that synaptic plasticity in recurrent excitatory to excitatory connections in ITC neurons alone can explain changes of activity patterns in both excitatory and inhibitory neurons with learning, it does not rule out synaptic plasticity in other synaptic connections of ITC neurons, or in other brain regions. Previous modeling work[14,39] implemented synaptic plasticity within recurrent circuits in order to explain effects of visual learning on delay activity during a memory task[40]. Other models have proposed learning at feedforward connections[41,42] or combined learning at feedforward and recurrent connections[15] to account qualitatively for the effect of familiarity on visual responses. None of these studies have quantitatively compared distributions of visual responses for novel and familiar stimuli, as we have done in our study. Also, note that our method could be generalized to feedforward or feedback circuits that have been suggested to be critical for object recognition[43–45], if the activities of input and output layers are given before and after learning. To disentangle learning occurring at multiple synaptic sites or in different areas, one could potentially use the different onset times of these synaptic inputs. For example, the effect of plasticity in feedforward connections would emerge immediately at the activity onset, while plasticity in the recurrent circuits would emerge with a delay corresponding to the time scales of intrinsic or synaptic time constants (Supplementary Fig. S7). Also, top-down signals can be affected by learning, whose effects would emerge with a longer latency of around 100 ms after the activity onset[46].

We explored extensions of the basic rule with a few schemes of synaptic weight normalization[34], and more general forms of learning rules incorporating *in vitro* experiments[6] and previous modeling work[47–50] need to be explored. Nevertheless, our method to infer learning rules from neuronal response distributions can provide a bridge between *in vitro* studies of synaptic plasticity rules and *in vivo* data obtained in behaving animals where synaptic changes are very difficult to measure directly, and could be applicable to other cortical circuits to further our understanding of the interactions between circuit dynamics and synaptic plasticity rules.

Online Methods

Behavioral task and neurophysiology

For investigation of visual experience in inferior temporal cortex (ITC), we compared visual responses to novel and familiar stimuli measured in different sets of monkeys (macaca mulatta) performing two different tasks (Fig. 1–4) in different laboratories. In a passive viewing task[22], monkeys fixated the stimuli that were presented for 200 ms with a 50 ms interval between stimuli. During the fixation, the responses of individual ITC neurons were measured with the stimulus sampled from 125 novel and 125 familiar images for each neuron. Activities of 73 putative excitatory neurons and 15 putative inhibitory neurons were recorded where cell types were classified from the width of their extra-cellularly recorded spikes[51] (for more details, see [22]).

In another experiment[25], monkeys performed a dimming-detection task, which is similar to the passive viewing task except that monkeys were required to detect and indicate (by releasing a manual lever) a subtle decrease in luminance of the stimulus. The purpose of the dimming task was to require the animal to direct their attention to the stimulus. 10 familiar and 8 novel images were presented during recordings for each neuron. Familiar images had been viewed many thousands of times over approximately one month of daily behavioral training sessions prior to recordings, and novel images had not been viewed before that session (but were repeated at least 10 times per session). Each dimming detection trial began when the animal grasped the manual lever, followed by the onset of a fixation spot. After acquiring gaze fixation (within a 2.0° radius window) for 500 ms, a single 100×100 pixel stimulus was presented foveally for a duration which is a sum of a fixed duration (650 ms) and a random duration (drawn from an exponential distribution with a mean of 200 ms). At the end of this duration, the imaged dimmed and the monkey was required to release the lever within 700 ms in order to receive a fruit juice reward. A total of 221 ITC cells were recorded from two monkeys, and as in the passive viewing task, putative neuronal classes were determined from the spike widths. While the distribution of spike widths showed two peaks, they were not far enough apart to separate two populations clearly. Thus, to minimize potential misclassifications, we set well separated thresholds, 500 μs and 250 μs for putative excitatory neurons and inhibitory neurons, respectively. With such thresholds, 41 and 27 neurons were classified as excitatory and inhibitory neurons, respectively.

The methods for neurophysiological recordings, surgery and behavioral task control for the second experiment have been described previously[52]. Briefly, neuronal recordings were conducted in both monkeys (male, 7–13 kg; research naïve prior to these experiments) from ITC (areas TEa, TEm, TE2 and TE1; [53]), including the cortex on the lower bank of the superior temporal sulcus and the ventral surface of the ITC, lateral to the anterior medial temporal sulcus. The recording chambers were surgically placed stereotaxically, guided by magnetic resonance imaging scans conducted prior to surgery. Neuronal recordings were conducted with 1–4 tungsten microelectrodes (100 μm diameter) using multiple motorized microdrives (NAN instruments) and multiple dura-piercing stainless steel guide tubes. All procedures were in accordance with the University of Chicago's Animal Care and Use Committee and US National Institutes of Health guidelines.

Data analysis in Figures 1 to 3

In both experiments, we obtained visual responses to each stimulus by taking the firing rate in the time window between 75 ms and 200 ms after stimulus onset. In Fig. 2 and 4 showing input changes and learning rules in individual neurons, 125 visual responses to novel and familiar stimuli measured in a passive fixation task were used. In Fig. 3 showing input changes from the distributions of firing rates averaged over many neurons, the number of visual responses was the product of the number of neurons and the number of stimuli (e.g., 88 neurons × 125 visual responses in Fig. 3a). In the analysis of learning effects on input currents in individual neurons (Fig. 4), we showed input changes only in neurons having significantly different distributions for novel and familiar stimuli, which was determined by the Mann-Whitney U test at 5% significance level (30 excitatory neurons in Fig. 4a and 10 inhibitory neurons in Fig. 4b). In Fig. 2g, Fig. 3 and Fig. 4a–d, noisy input changes were

smoothed for presentation – after interpolating input changes at the equally spaced firing rates, we smoothed input currents locally using lowess (Locally Weighted Scatterplot Smoothing) function with a span of 10% in Matlab. The significance of the input changes was computed by sampling multiple times from the same distribution of novel stimuli and computing the resulting input changes. We show a 95% confidence level, that is, ± 1.96 times the standard deviation of input changes at each post-synaptic firing rate.

Network model

To reproduce activity changes observed in the data, we considered a firing rate model and a rate-based plasticity rule in the excitatory to excitatory connections. The network was composed of N_E excitatory and N_I inhibitory neurons,. The firing rate of each neuron is denoted by r_i^l with the superscript $l = E$ or I, and the subscript i represents the index of the neuron, ranging between 1 and N_l. These firing rates are governed by the equations

$$\begin{aligned}\tau_E \frac{dr_i^E}{dt} &= -r_i^E + \Phi_E\left(\sum_{j=1}^{N_E} W_{ij}^{EE} r_j^E - \sum_{j=1}^{N_I} W_{ij}^{EI} r_j^I + I_i^{EX}\right)\\ \tau_I \frac{dr_i^I}{dt} &= -r_i^I + \Phi_I\left(\sum_{j=1}^{N_E} W_{ij}^{IE} r_j^E + I_i^{IX}\right)\end{aligned}$$

where W_{ij}^{lm} is the strength of the synaptic connection from neuron j in population m to neuron i in population l with l, or $m = E$, or I, $i = 1$ to N_l, and $j = 1$ to N_m. Thus, excitatory neuron i receives recurrent inputs from excitatory and inhibitory populations and the external input I_i^{EX}, while inhibitory neurons receive recurrent excitatory inputs and the external input I_i^{EX}. The firing rate r_i^l approaches $\Phi_l\left(x_i^l\right)$ with intrinsic time constant τ_l, where $\Phi_l\left(x_i^l\right)$ represents the steady state firing rate in response to total input current x_i^l.

We assumed only the excitatory to excitatory connections are plastic, and the amount of the synaptic change depends on firing rates of pre- and post-synaptic neurons. We assumed that the dependence on pre- and post-synaptic terms is separable as $\Delta W_{ij}^{EE} = f_{post}\left(r_i^E\right) f_{pre}\left(r_j^E\right)$. Furthermore, we assumed that $f_{pre}\left(r_j^E\right) = r_j^E$, and to prevent too low mean weights after learning multiple stimuli, we considered a constraint that the sum of synaptic weights over the pre-synaptic neurons is preserved with learning as

$$\begin{aligned}\Delta W_{ij}^{EE} \leftarrow\ & \Delta W_{ij}^{EE} - \frac{1}{N_E}\sum_{j=1}^{N_E} \Delta W_{ij}^{EE}\\ =\ & f_{post}\left(r_i^E\right) r_j^E - \frac{1}{N_E}\sum_{j=1}^{N_E} f_{post}\left(r_i^E\right) r_j^E\\ =\ & f_{post}\left(r_i^E\right)\left(r_j^E - mean\left(r^E\right)\right),\end{aligned}$$

with W_{ij}^{EE} having lower bound 0 and upper bound w_{EE}^{max}/N_E. Note that this is equivalent to taking $f_{pre}\left(r_j^E\right) = r_j^E - mean\left(r^E\right)$. The remaining synaptic connectivity was assumed to be uniform, that is $W_{ij}^{EI} = w_{EI}/N_I$ and $W_{ij}^{IE} = w_{IE}/N_E$ for all indices *i* and *j*.

In the simulation, N_E = 4000, N_I = 1000, τ_E = 20ms and τ_I = 10ms. w_{EE}^{max}=0.1, w_{EI}=0.01 and w_{IE}=0.5, with the initial W_{ij}^{EE} set to be $w_{EE}^{max}/2N_E$. The activity pattern for novel stimuli, input current-output firing rate transfer function Φ_i, and the dependence of learning rules on post-synaptic firing rates $f_{post}\left(r_i^E\right)$ were obtained from the individual neuronal responses for novel and familiar stimuli averaged over neurons showing characteristic input changes, that is, excitatory neurons showing both depression and potentiation (green curves in Fig. 4a) and inhibitory neurons showing only depression (orange curves in Fig. 4b). In particular, $f_{post}\left(r_i^E\right)$ was derived from the dependence of input changes on post-synaptic firing rates Δh_i as

$$\begin{aligned} f_{post}(r_i) &\approx \left(\Delta h_i - \sum_{j=1}^{N_E} W_{ij}^{EE}\Delta r_j^E + \sum_{j=1}^{N_I} W_{ij}^{EI}\Delta r_j^I\right) / \sum_j f_{pre}(r_j)\, r_j \\ &\approx \frac{\Delta h_i - N_E m\left(W_{ij}^{EE}\right)\left(m\left(r^{E,fam}\right) - m\left(r^{E,nov}\right)\right) + w_{EI}\left(m\left(r^{I,fam}\right) - m\left(r^{I,nov}\right)\right)}{N_E\, var\left(r^{E,nov}\right)} \end{aligned}$$

Where $m(r^{I,fam})$ and $m(r^{I,nov})$ are average firing rates for familiar and novel stimuli of population $I = E$ or I, and var($r^{E,nov}$) is the variance of the firing rates for novel stimuli, obtained from the data. $m\left(W_{ij}^{EE}\right)$ is the average of excitatory to excitatory synaptic weights and it is updated during initialization of the excitatory connections as the network learns multiple uncorrelated activity patterns. Due to the constraint on the sum of the synaptic weights, $N_E m\left(W_{ij}^{EE}\right) \sim w_{EE}^{max}/2$, that is, it remains close to its initial value.

Code availability

The data analysis and network simulations were performed in Matlab, and the code for network simulations will be available upon request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank S. Dieudonné, and D. Higgins for valuable discussions and Y. Aljadeff, K. Burbank, and M. de Pittà for feedback on the manuscript. D.L.S. has been supported by grants from the National Science Foundation (SBE-0542013) and the US National Institutes of Health (R01EY14681). D.J.F is a recipient of NSF CAREER Award, McKnight Scholar award, and has also been supported by the Alfred P. Sloan Foundation, and a Natural Sciences and Engineering. J.L.M is a recipient of Research Council of Canada (NSERC) fellowship.

References

1. Hebb, DO. The Organization of Behavior. Wiley; 1949.
2. Bi GQ, Poo MM. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J. Neurosci. 1998; 18:10464–10472. [PubMed: 9852584]
3. Bliss TV, Lomo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J. Physiol. 1973; 232:331–356. [PubMed: 4727084]
4. Dudek SM, Bear MF. Homosynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. Proc. Natl. Acad. Sci. U. S. A. 1992; 89:4363–4367. [PubMed: 1350090]
5. Markram H, Lubke J, Frotscher M, Sakmann B. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science. 1997; 275:213–215. [PubMed: 8985014]
6. Sjostrom PJ, Turrigiano GG, Nelson SB. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron. 2001; 32:1149–1164. [PubMed: 11754844]
7. Dan Y, Poo MM. Spike timing-dependent plasticity: from synapse to perception. Physiol. Rev. 2006; 86:1033–1048. [PubMed: 16816145]
8. Feldman DE. Synaptic mechanisms for plasticity in neocortex. Annu. Rev. Neurosci. 2009; 32:33–55. [PubMed: 19400721]
9. Fox K, Wong RO. A comparison of experience-dependent plasticity in the visual and somatosensory systems. Neuron. 2005; 48:465–477. [PubMed: 16269363]
10. Dayan, P.; Abbott, LF. Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. The MIT Press; 2005.
11. Gerstner, W.; Kistler, WM. Spiking Neuron Models: Single Neurons, Populations, Plasticity. Cambridge University Press; 2002.
12. Bienenstock EL, Cooper LN, Munro PW. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J. Neurosci. 1982; 2:32–48. [PubMed: 7054394]
13. Song S, Miller KD, Abbott LF. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nat. Neurosci. 2000; 3:919–926. [PubMed: 10966623]
14. Amit DJ, Brunel N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb. Cortex. 1997; 7:237–252. [PubMed: 9143444]
15. Sohal VS, Hasselmo ME. A model for experience-dependent changes in the responses of inferotemporal neurons. Network. 2000; 11:169–190. [PubMed: 11014667]
16. Miyashita Y. Inferior temporal cortex: where visual perception meets memory. Annu. Rev. Neurosci. 1993; 16:245–263. [PubMed: 8460893]
17. Tanaka K. Inferotemporal cortex and object vision. Annu. Rev. Neurosci. 1996; 19:109–139. [PubMed: 8833438]
18. Kobatake E, Wang G, Tanaka K. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. J. Neurophysiol. 1998; 80:324–330. [PubMed: 9658053]
19. Li L, Miller EK, Desimone R. The representation of stimulus familiarity in anterior inferior temporal cortex. J. Neurophysiol. 1993; 69:1918–1929. [PubMed: 8350131]
20. Logothetis NK, Pauls J, Poggio T. Shape representation in the inferior temporal cortex of monkeys. Curr. Biol. 1995; 5:552–563. [PubMed: 7583105]
21. Freedman DJ, Riesenhuber M, Poggio T, Miller EK. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. Cereb. Cortex. 2006; 16:1631–1644. [PubMed: 16400159]
22. Woloszyn L, Sheinberg DL. Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex. Neuron. 2012; 74:193–205. [PubMed: 22500640]
23. Meyer T, Walker C, Cho RY, Olson CR. Image familiarization sharpens response dynamics of neurons in inferotemporal cortex. Nat. Neurosci. 2014; 17:1388–1394. [PubMed: 25151263]

24. Op de Beeck HP, Wagemans J, Vogels R. Effects of perceptual learning in visual backward masking on the responses of macaque inferior temporal neurons. Neuroscience. 2007; 145:775–789. [PubMed: 17293053]
25. McKee, JL.; Thomas, SL.; Freedman, DJ. Society for Neuroscience. San Diego: 2013. Abstracts
26. Buzsaki G, Mizuseki K. The log-dynamic brain: how skewed distributions affect network operations. Nat. Rev. Neurosci. 2014; 15:264–278. [PubMed: 24569488]
27. Hromadka T, Deweese MR, Zador AM. Sparse representation of sounds in the unanesthetized auditory cortex. PLoS Biol. 2008; 6:e16. [PubMed: 18232737]
28. Roxin A, Brunel N, Hansel D, Mongillo G, van Vreeswijk C. On the distribution of firing rates in networks of cortical neurons. J. Neurosci. 2011; 31:16217–16226. [PubMed: 22072673]
29. Anderson JS, Lampl I, Gillespie DC, Ferster D. The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. Science. 2000; 290:1968–1972. [PubMed: 11110664]
30. Rauch A, La Camera G, Luscher HR, Senn W, Fusi S. Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo-like input currents. J. Neurophysiol. 2003; 90:1598–1612. [PubMed: 12750422]
31. Kirkwood A, Rioult MC, Bear MF. Experience-dependent modification of synaptic plasticity in visual cortex. Nature. 1996; 381:526–528. [PubMed: 8632826]
32. Toyoizumi T, Kaneko M, Stryker MP, Miller KD. Modeling the dynamic interaction of hebbian and homeostatic plasticity. Neuron. 2014; 84:497–510. [PubMed: 25374364]
33. Zenke F, Hennequin G, Gerstner W. Synaptic plasticity in neural networks needs homeostasis with a fast rate detector. PLoS Comput. Biol. 2013; 9:e1003330. [PubMed: 24244138]
34. Miller KD, MacKay DJ. C. The Role of Constraints in Hebbian Learning. Neural Comput. 1994; 6:100–126.
35. Bourne JN, Harris KM. Coordination of size and number of excitatory and inhibitory synapses results in a balanced structural plasticity along mature hippocampal CA1 dendrites during LTP. Hippocampus. 2011; 21:354–373. [PubMed: 20101601]
36. Amit DJ, Fusi S. Learning in neural networks with material synapses. Neural Comput. 1994; 6
37. Sejnowski TJ. Storing covariance with nonlinearly interacting neurons. J. Math. Biol. 1977; 4:303–321. [PubMed: 925522]
38. Cooper, LN.; Intrator, N.; Blais, BS.; Shouval, HZ. Theory of cortical plasticity. World Scientific Publishing Co. Pte. Ltd.; 2004.
39. Brunel N. Hebbian learning of context in recurrent neural networks. Neural Comput. 1996; 8:1677–1710. [PubMed: 8888613]
40. Miyashita Y. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. Nature. 1988; 335:817–820. [PubMed: 3185711]
41. Bogacz R, Brown MW. Comparison of computational models of familiarity discrimination in the perirhinal cortex. Hippocampus. 2003; 13:494–524. [PubMed: 12836918]
42. Norman KA, O'Reilly RC. Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. Psychol. Rev. 2003; 110:611–646. [PubMed: 14599236]
43. Riesenhuber M, Poggio T. Hierarchical models of object recognition in cortex. Nat. Neurosci. 1999; 2:1019–1025. [PubMed: 10526343]
44. Yamins DL, et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc. Natl. Acad. Sci. U. S. A. 2014; 111:8619–8624. [PubMed: 24812127]
45. DiCarlo JJ, Zoccolan D, Rust NC. How does the brain solve visual object recognition? Neuron. 2012; 73:415–434. [PubMed: 22325196]
46. Tomita H, Ohbayashi M, Nakahara K, Hasegawa I, Miyashita Y. Top-down signal from prefrontal cortex in executive control of memory retrieval. Nature. 1999; 401:699–703. [PubMed: 10537108]
47. Clopath C, Busing L, Vasilaki E, Gerstner W. Connectivity reflects coding: a model of voltage-based STDP with homeostasis. Nat. Neurosci. 2010; 13:344–352. [PubMed: 20098420]
48. Graupner M, Brunel N. Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:3991–3996. [PubMed: 22357758]

49. Pfister JP, Gerstner W. Triplets of spikes in a model of spike timing-dependent plasticity. J. Neurosci. 2006; 26:9673–9682. [PubMed: 16988038]

50. Shouval HZ, Bear MF, Cooper LN. A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. Proc. Natl. Acad. Sci. U. S. A. 2002; 99:10831–10836. [PubMed: 12136127]

51. McCormick DA, Connors BW, Lighthall JW, Prince DA. Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. J. Neurophysiol. 1985; 54:782–806. [PubMed: 2999347]

52. Ibos G, Freedman DJ. Dynamic integration of task-relevant visual features in posterior parietal cortex. Neuron. 2014; 83:1468–1480. [PubMed: 25199703]

53. Paxinos, G.; Huang, XF.; Toga, AW. The rhesus monkey brain in stereotaxic coordinates. Academic Press; 2000.

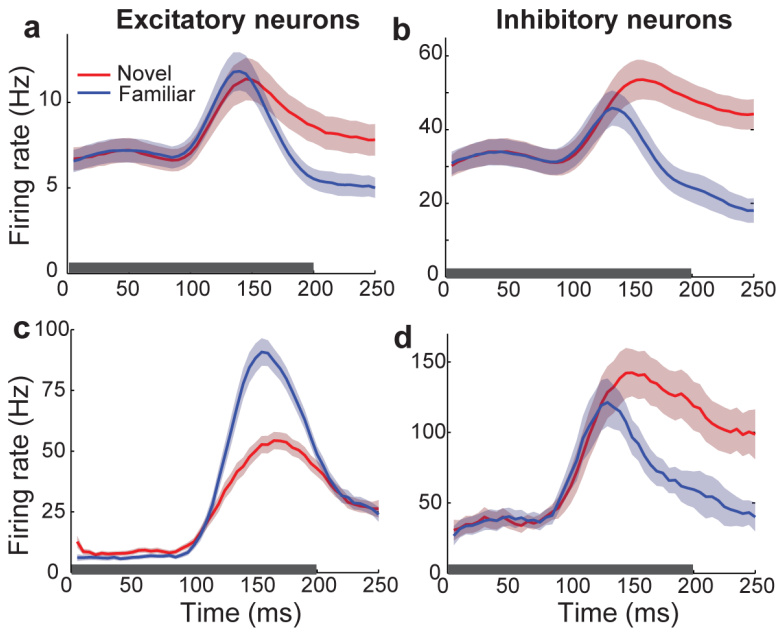

Fig. 1. Visual response of inferior temporal cortical (ITC) neurons to novel and familiar stimuli

a–d, Time course of mean (**a,b**) and maximal (**c,d**) visual responses of ITC excitatory (**a,c**) and inhibitory (**b,d**) neurons obtained in a passive viewing task[22]. Solid curves are activities averaged over all neuron for all novel (red) or familiar (blue) stimuli, and error bars represent mean ± SEM of activities averaged over individual neurons (**a–d**). The grey horizontal bar represents the visual stimulation period. For more details of the experiment, see Online Methods and [22].

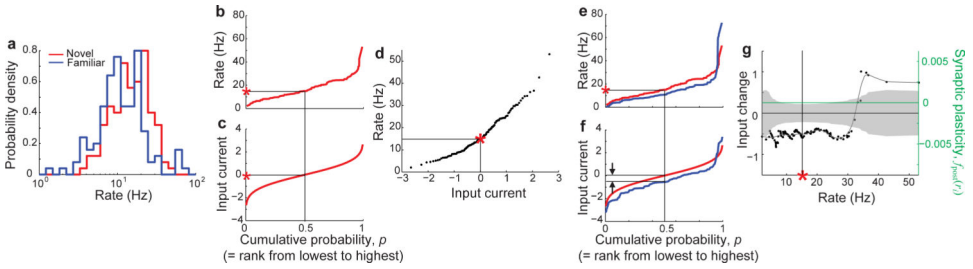

Fig. 2. Inferring learning rules from distributions of firing rates

a, Distributions of firing rates of a single ITC neuron in response to novel (red) and familiar stimuli (blue). **b–d,** Deriving a static transfer function Φ from the distribution of visual responses for novel stimuli. **b,c,** Inverse cumulative distribution functions of firing rates (**b**) and input current (**c**) to novel stimuli, which can be obtained by ordering rates and input currents according to their rank from lowest to highest. Input currents were assumed to follow Gaussian statistics and were normalized by their means and standard deviations yielding mean and variance 0 and 1, respectively. Red asterisks are the median firing rate and input current ($p = 0.5$). **d,** Input current-output firing rate transfer function Φ. The red asterisk in **d** shows the firing rate and input current for $p = 0.5$ (see red asterisks in **b** and **c**). **e–g**, Inferring input changes and learning rules. **e,f,** Inverse cumulative distribution functions of firing rates (**e**) and input currents (**f**) for novel (red) and familiar (blue) stimuli. **g**, Input changes and dependence of learning rule on the post-synaptic firing rate. The dependence of synaptic plasticity rule on the post-synaptic firing rate is similar to input changes with a constant offset and rescaling of its magnitude, where the constant offset and scaling factor are chosen to be 0.25 and 200, respectively, for illustration (the same data points with green axis on the right). Grey area represents a 95% confidence region (see Online Methods).

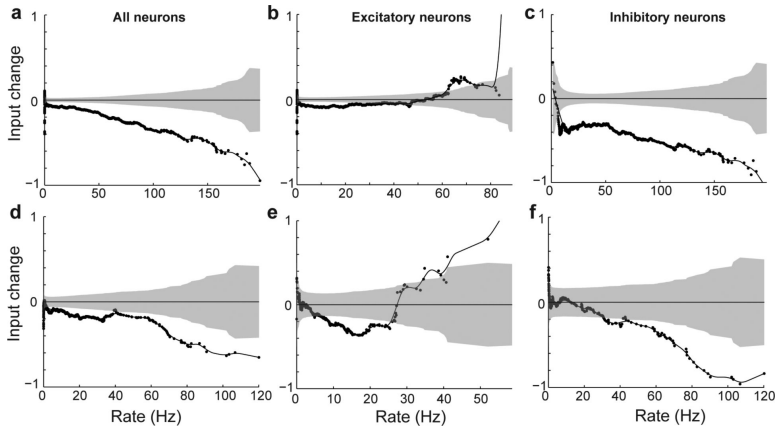

Fig. 3. Effect of visual experience in ITC neurons and their dependence on different cell types

a–c, Input changes obtained from visual responses of ITC neurons in monkeys performing a passive viewing task. The input currents were obtained from the distributions of firing rates for novel and familiar stimuli that were averaged over all recorded neurons (n = 88). Only negative changes are observed with learning (**a**, see Online Methods for smoothing procedure). When neurons were grouped as putative excitatory (n = 73) and inhibitory neurons (n = 15), excitatory neurons showed negative changes for low post-synaptic firing rate, and positive changes for high rate (**b**), while inhibitory neurons showed only negative changes (**c**). **d–f,** Input changes obtained in monkeys performing a dimming-detection task. As in a passive viewing task (**a–c**), the input changes obtained from all recorded neurons (**d**, n = 221) and from putative inhibitory neurons (**f**) showed only decrease, while putative excitatory neurons showed both negative and positive changes (**e**). Note that putative neuronal classes were determined from the spike widths. To minimize potential misclassifications in the dimming-detection task, we set well separated thresholds for putative excitatory and inhibitory neurons, leading to 41 and 27 neurons classified as excitatory and inhibitory neurons, respectively (see Online methods).

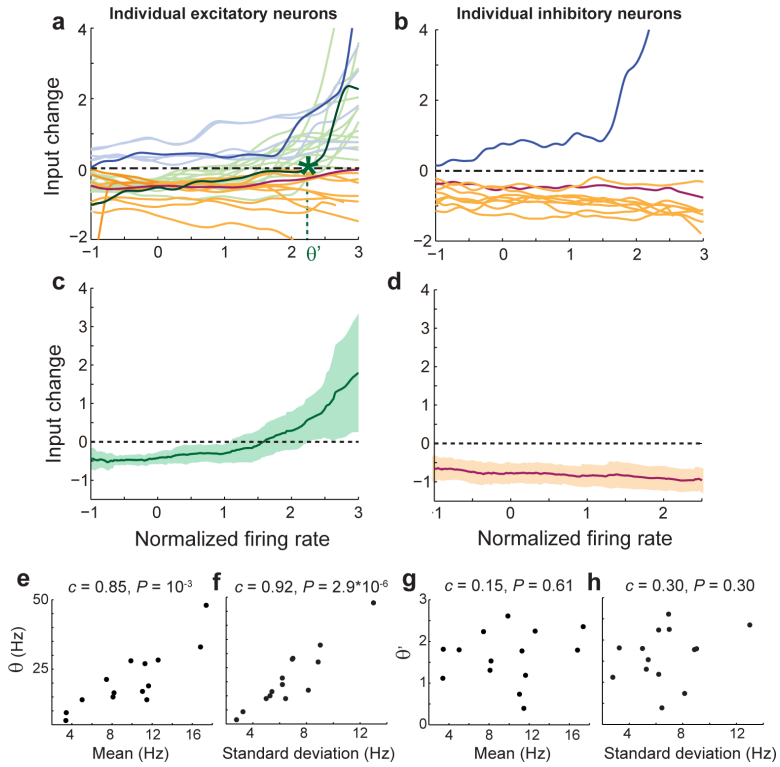

Fig. 4. Effect of visual experience in individual ITC neurons and regulation of learning rules

a,b, Input changes obtained from visual responses of individual excitatory (**a**, n = 30) and inhibitory (**b**, n = 10) neurons. Firing rates of post-synaptic neurons to novel stimuli were normalized with their mean and standard deviation, and input changes were smoothed (see Online Methods). Neurons were classified into 3 categories as neurons showing only negative changes (orange; 10 excitatory and 9 inhibitory neurons), only positive changes (blue; 6 excitatory and 1 inhibitory neurons), and both negative and positive changes (green; 14 excitatory neurons). Dark-colored curves are example neurons in each class. **c,d,** Average input changes in excitatory neurons (**c**) and inhibitory neurons (**d**) with colored areas representing variabilities over different neurons. Average input changes in excitatory neurons (**c**) and inhibitory neurons (**d**) were obtained from neurons showing negative changes for low rates and positive changes for high rates (green curves in **a**) and neurons showing only negative changes (orange curves in **b**), respectively. **e,h,** Correlation between activity of neurons and a threshold firing rate separating potentiation from depression in input currents. For excitatory neurons showing both negative and positive changes (green curves in **a**), a threshold θ was defined as the firing rate for which input changes become positive, and θ' was a normalized θ with the mean and standard deviation of post-synaptic firing rates (green asterisk in **a**). *c* and *P* in each plot are the correlation coefficients and their *P*-values.

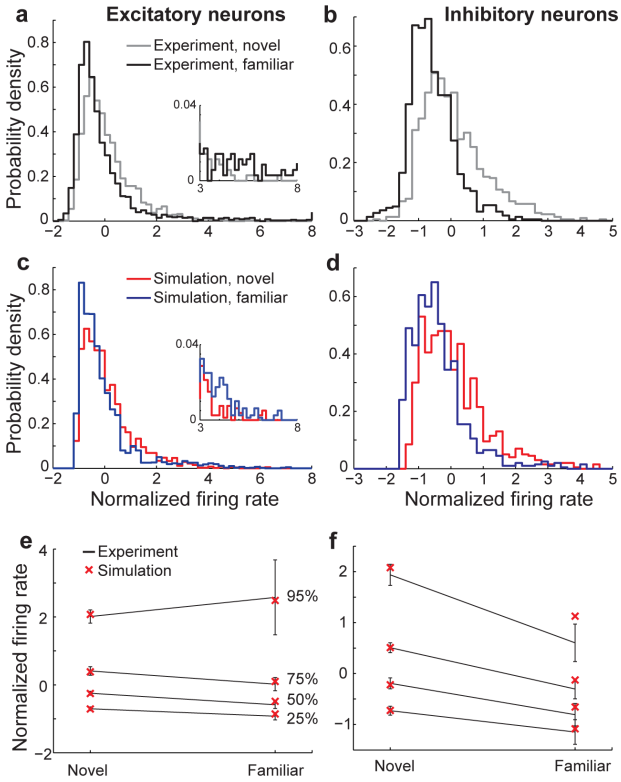

Fig. 5. Comparison between simulated and experimental data

a,b, Distributions of normalized firing rates of excitatory (**a**) and inhibitory (**b**) neurons for novel (grey) and familiar (black) stimuli, obtained from the experiment. The distributions were obtained from the activities of individual neurons showing characteristic input changes in each cell type (green curves in Fig. 4a for excitatory neurons and orange curves in Fig. 4b for inhibitory neurons) – firing rates for novel stimuli were normalized by the mean and standard deviation for each individual neuron and distributions of normalized firing rates were averaged over neurons in each cell type (grey). The distributions for familiar stimuli (black) were obtained similarly, except that firing rates were normalized with the mean and standard deviation of firing rates to novel stimuli to trace changes with learning. **c,d,** Distributions of normalized firing rates for novel (red) and familiar (blue) stimuli, obtained from the simulation. As in the data, the firing rates were normalized with mean and standard deviation of firing rates for novel stimuli. **e,f,** Changes of 25th, 50th, 75th and 95th percentiles of normalized firing rates in the data (black) and in the simulation (red cross). In the data, we averaged percentiles obtained in individual neurons, and error bar is the standard deviation of normalized firing rates over different neurons.